high availability
systems
engineering

july 2003

technical
white paper

HP High Performance Clusters LC Series setup and installation guide

table of contents

341524- 001

abstract

This setup and installation guide provides the customer with the information required to take the HP High Performance Clusters LC Series (HPC LC Series) product offering from delivery of order to a fully cabled system, ready for operating system and application installation. This document supplements the information found in the user guides for the servers and switches used in your HPC LC Series solution.

introduction

This guide provides the customer with instructions and the necessary reference information required for setting up and installing the HP High Performance Clusters LC Series solutions. The reader gains the knowledge of the various solution components and how to receive equipment, position racks, cable the configuration, power on the equipment, and begin the initial setup process. This paper provides information on each of the following topics:

- HPC LC Series hardware solutions
- HPC LC Series solution components
- HP Services and Care Packs
- Receiving the HPC cluster
- Installing switches
- Physical Planning
- Cabling the configuration
- Powering on the equipment
- Default factory settings
- Setting up the control node
- Setting up the compute nodes

This paper assumes that the reader has advanced technical skills and has extensive knowledge of the following topics and products:

- High performance computing concepts
- Linux operating system installation knowledge and experience
- High performance computing system software installation experience

safety information

⚠ IMPORTANT SAFETY INFORMATION

Before installation, read the *Important Safety Information* document included with the product. Also, read the safety information details of the documentation included for each component.

HPC LC Series hardware solutions

The HPC LC Series is a range of pre-configured hardware configurations that are integrated, tested, and shipped assembled in racks ready for customer use. The solutions utilize the ProLiant DL380 G3 server as the control node (service node) and the ProLiant DL360 G3 servers as the compute nodes (application nodes). The customer simply selects the appropriate HPC LC Series configuration that represents the required number of compute nodes and the desired speed of the cluster interconnect.

There are four (4) cluster node size categories:

- 16-node
- 32-node
- 64-node
- 128-node

There are three (3) cluster interconnect categories:
- 10/100 Fast Ethernet
- Gigabit Ethernet
- Myrinet

For each region, there are a total of eleven (11) HPC LC Series solutions to choose from based on possible combinations of the cluster node size and the type of cluster interconnect.

Table 1 below lists all eleven part numbers for North America. Table 2 below lists all eleven part numbers for Europe, Middle East, and Africa (EMEA).

table 1. North America: HPC LC Series hardware solutions

North America		Cluster Node Size			
		16-Node	**32-Node**	**64-Node**	**128-Node**
Cluster interconnect	**10/100 Fast Ethernet**	Part number: 322939-001	Part number: 322940-001	Part number: 322941-001	Part number: 322942-001
	Gigabit Ethernet	Part number: 322939-002	Part number: 322940-002	Part number: 322941-002	Not offered at this time
	Myrinet switch	Part number: 322939-003	Part number: 322940-003	Part number: 322941-003	Part number: 322942-003

table 2. EMEA: HPC LC Series hardware solutions

EMEA		Cluster Node Size			
		16-Node	**32-Node**	**64-Node**	**128-Node**
Cluster interconnect	**10/100 Fast Ethernet**	Part number: 322939-421	Part number: 322940-421	Part number: 322941-421	Part number: 322942-421
	Gigabit Ethernet	Part number: 322939-422	Part number: 322940-422	Part number: 322941-422	Not offered at this time
	Myrinet switch	Part number: 322939-423	Part number: 322940-423	Part number: 322941-423	Part number: 322942-423

HPC LC Series solution components

Each HPC LC Series cluster contains one control node, several compute nodes, interconnects, rack(s), and rack infrastructure. The next few sections describe each of these components in more detail.

control node

One ProLiant DL380 G3 server functions as the control node (service node), and is equipped with one Intel Pentium 4 - 3.06 GHz processor, 1 GB DDR memory, and one 36 GB Ultra 320 10K SCSI Disk as a base node configuration. This server also comes with a redundant power supply, a redundant fan kit, and a dual port PCI Gigabit NIC adapter. The control node can be extended up to two processors, twelve GB DDR memory, and six Ultra 320 SCSI hard drives. The server may also have a PCI Myrinet adapter installed if this server is part of a HPC Myrinet solution. This control node is used as the interface to the user community for job dispatch, control, monitoring, and job completion within the cluster.

compute nodes

Depending on which configuration you ordered, you would have from 15 to 127 ProLiant DL360 G3 servers that function as compute nodes (application nodes). Each compute node is equipped with one Intel Pentium 4 - 3.06 GHz processor and 1 GB DDR memory. The compute node systems can be extended up to two processors, eight GB DDR memory, and two SCSI hard drives. Each compute node may also have a PCI Myrinet adapter installed if the node is part of a HPC Myrinet solution.

HPC networks

Each HPC LC Series cluster is made up of multiple networks which include the following: iLO network, cluster interconnect network, and management network. Each one of these networks is described in more detail below.

- <u>iLO network (Out of Band Management)</u> – This network is used for integrated Lights-Out (iLO) connections to all of the ProLiant servers that make up the HPC configuration. This network uses the HP ProCurve 2650 switch.
- <u>cluster interconnect network</u> – This is the main data network that connects all of the compute nodes together. This network can be either a 10/100 Fast Ethernet, Gigabit Ethernet, or Myrinet network. This network utilizes various types of ProCurve or Myrinet switches depending on the overall type and size of the HPC solution.
- <u>management network (In Band Management)</u> – This network is used for overall cluster management using a standard Ethernet connection. If using the 10/100 Fast Ethernet or Gigabit Ethernet HPC solutions, you have the option to have the management network on a separate hardware switch. If using the Myrinet HPC solutions, the management network is on a separate hardware switch. This network uses the HP ProCurve 2650 switch.

additional components

Each HPC LC Series solution also comes equipped with an HP TFT 5600 RKM, Modular 24A PDU(s), and extra network cables for external network connectivity. All components are integrated and pre-cabled into the 42U rack(s). All of the internal rack network cables are labeled with a descriptive cable label to facilitate the identification process of each cable connection. The *HP High Performance Clusters LC Series Cabling Guide* illustrates the point-to-point connections of each network cable and describes the cable label nomenclature in detail.

hp services and care packs

To facilitate the setup and installation process of your HPC LC Series solution, HP offers a variety of installation and professional services. HP's highly skilled professionals can help you manage information resources, provide consistent performance, and deliver secure access with our comprehensive suite of proactive services. HP Care Pack services offer upgraded service levels to extend and expand your standard product warranty with easy-to-buy and easy-to-use support packages that help you make the most of your hardware and software investments.

To learn more about what services are available for your HPC LC Series solution, please visit the following:

- hp services: http://www.hp.com/hps/
- hp care pack services: http://www.hp.com/hps/carepack

receiving the HPC cluster

Since the HPC LC Series cluster sizes range from 16-nodes to 128-nodes, the cluster size will determine if one, two, or four 42U racks are delivered. The 16-node and 32-node solutions ship in one 42U rack. The 64-node solutions ships in two 42U racks. The 128-node solutions ship in four 42U racks. Every configuration is shipped fully integrated with easy to read cable labels to facilitate the cabling process. Depending on which HPC LC Series solution that was ordered, the cluster interconnect switch may be shipped separately instead of being integrated and shipped in the rack. It is necessary to ship some of the cluster interconnect switches separately to prevent damage to the unit during the shipping process. Refer to the next section for which HPC LC Series solutions may require a switch to be shipped separately and installed at the customer site.

installing switches

Some of the HPC LC Series solutions may contain a switch that cannot be shipped in the rack. In these cases, the switch is shipped in a separate box or boxes. The following HPC solutions may contain a switch that is shipped separately:

- 16-Node Myrinet Solution - the Myrinet (3-slot frame) switch may ship separately*
- 32-Node Myrinet Solution - the Myrinet (5-slot frame) switch may ship separately*
- 64-Node Myrinet Solution - the Myrinet (9-slot frame) switch may ship separately*
- 64-Node Gigabit Solution - the ProCurve 9308m switch may ship separately*
- 128-Node Myrinet Solution - the Myrinet (17-slot frame) switch may ship separately*

* This solution may contain a switch that is shipped separately. If you received a switch that was shipped separately, then the switch must be installed in the pre-defined space within the rack at the customer site. Refer to the *HP High Performance Clusters LC Series Cabling Guide* for information on switch placement within the rack. In the future, the switch may be shipped in the rack, instead of being shipped separately.

physical planning

Physical planning for your HPC LC Series deployment is one of the first things that must be considered before beginning the installation. You must ensure that you have enough physical space, adequate power and ventilation. You should also provide backup power such as an Uninterruptible Power Supply (UPS). A properly designed computer room has adequate ventilation and cooling for racks with servers and storage devices and has the appropriate high-line power feeds installed. For more information on datacenter design and planning, please refer to Technology Brief TC030203TB at the link below. This technology brief describes trends affecting datacenter design, explains how to determine power and cooling needs, and describes methods for cost-effective cooling.

Technology Brief TC030203TB can be downloaded from the following link:
http://wwss1pro.compaq.com/support/reference_library/viewdocument.asp?countryco de=1000&prodid=137&source=tc030203tb.xml&dt=21&docid=15719

positioning the racks

Upon receipt of the HPC LC Series solution, the racks will need to be placed in the appropriate positions within the customer data center. The 16-node solutions ship in a single 42U rack and require sufficient power for two (2) Power Distribution Units (PDUs). The 32-node solutions also ship in a single 42U rack and require sufficient power for three (3) PDUs. The 64-node solutions ship in two 42U racks and require sufficient power for six (6) PDUs. The 128-node solutions ship in four racks and require sufficient power for eleven (11) PDUs. When positioning the racks, be sure to place them in sequential order because the cable lengths are designed for that relative position. For example, figure 1 below illustrates the proper rack positioning of a 128-node cluster solution.
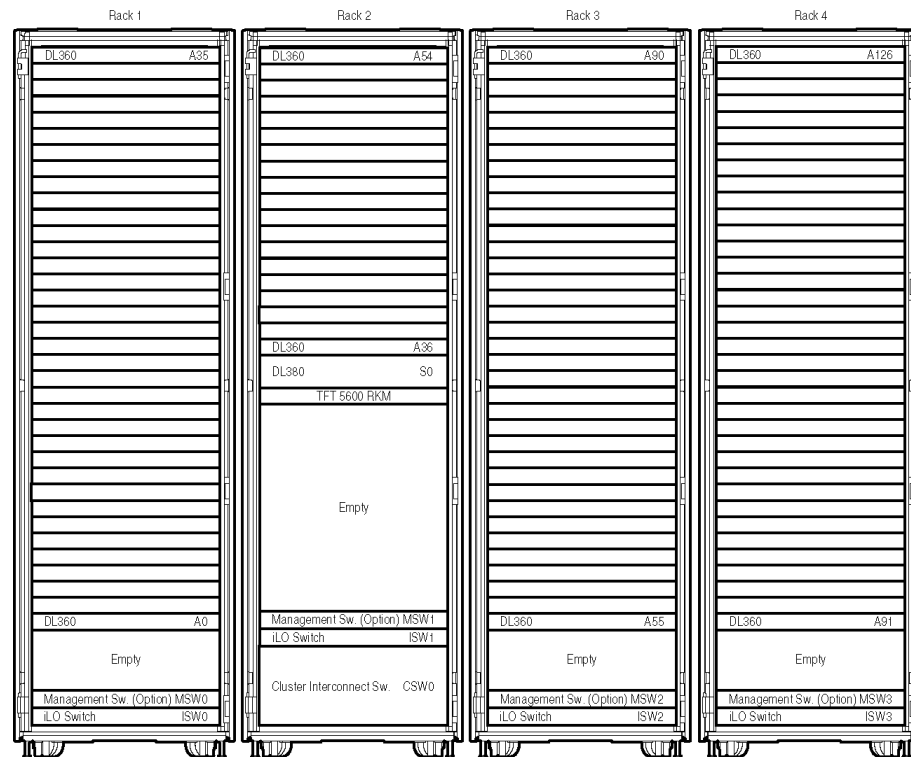

figure 1. rack positioning for a HPC LC Series solution

The following spatial needs should be considered when deciding where to physically place the HPC LC Series cluster solutions:

- Clearance in front of the rack unit should be a minimum of 25 inches for the front doors to open completely and for adequate airflow.

- Clearance behind the rack unit should be a minimum of 30 inches to allow for servicing and for adequate airflow.

rack cooling

The racks in each HPC LC Series solution draw cool air in through the front and exhaust warm air out of the rear. To ensure continued safe and reliable operation of the equipment, place the system in a well-ventilated, climate-controlled environment. The HPC LC Series solutions should be placed in data centers with an adequate air-conditioning system to handle continuous operation of this solution. To prevent component overheating and thermal shutdowns, please review the documentation for each of the components within your HPC LC Series solution to learn more about the recommended ambient (inlet) temperatures and the allowable maximum ambient operating temperatures within your data center.

cabling the configuration

All of the network cables within each rack are labeled for easy identification. The *HP High Performance Clusters LC Series Cabling Guide* explains and illustrates the cabling requirements for each of the HPC LC Series solutions in detail. Also, if your HPC solution is comprised of multiple racks, then there will be some inter-rack cabling. That is, some of the cables from one rack will be connected to a switch in another rack. Furthermore, each HPC solution will have some network cables that need to be connected outside of the rack to the corporate network (for example, the DHCP connection). The network cables that need to be connected outside of the rack are not labeled but are included with the product. Refer to figures 2 and 3 for a high level overview of the network cabling requirements for the various HPC LC Series solutions.

IMPORTANT: Refer to the *HP High Performance Clusters LC Series Cabling Guide* for details on the power and network cabling requirements, cable label nomenclature, and cabling illustrations for each HPC LC Series solution.

Figure 2 below provides a high level overview of the network cabling requirements for the **10/100 Fast Ethernet** and **Gigabit Ethernet** HPC cluster solutions only.
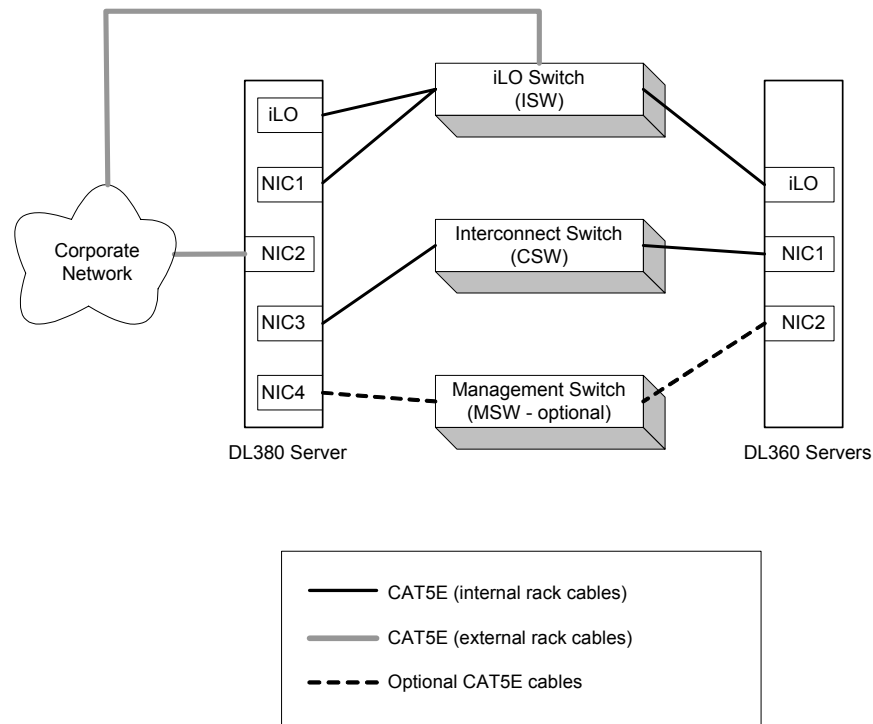

figure 2. network overview of 10/100 Fast Ethernet and Gigabit Ethernet HPC solutions

In figure 2 above, the DL380 server has one (1) iLO port and four (4) NIC ports. Both the iLO port and NIC1 port connect to the iLO switch (ISW). NIC2 is used to connect to an external network. NIC3 is used to connect to the cluster interconnect switch (CSW). In this configuration, the cluster interconnect switch can be either a 10/100 Fast Ethernet or Gigabit Ethernet switch. NIC4 is used to connect to the optional management switch (MSW). If the management switch is not part of the solution, then NIC4 will not be connected.

The DL360 servers have one (1) iLO port and two (2) NIC ports. The iLO port connects to the iLO switch (ISW). NIC1 connects to the cluster interconnect switch (CSW). NIC2 connects to the optional management switch (MSW). If the management switch is not part of the solution, then NIC2 will not be connected.

Note: The management switch (MSW) is an optional component for the 10/100 Fast Ethernet and Gigabit Ethernet solutions.

Figure 2 above assumes you will be connecting your HPC LC Series solution to an external DHCP server. The DHCP server must be provided by the customer and be made available to the system before proceeding to use iLO. If you plan on using the control node and HPC cluster software to assign DHCP addresses, then you will need to cable the system accordingly.

Figure 3 below provides a high level overview of the network cabling requirements for the **Myrinet** HPC cluster solutions only.
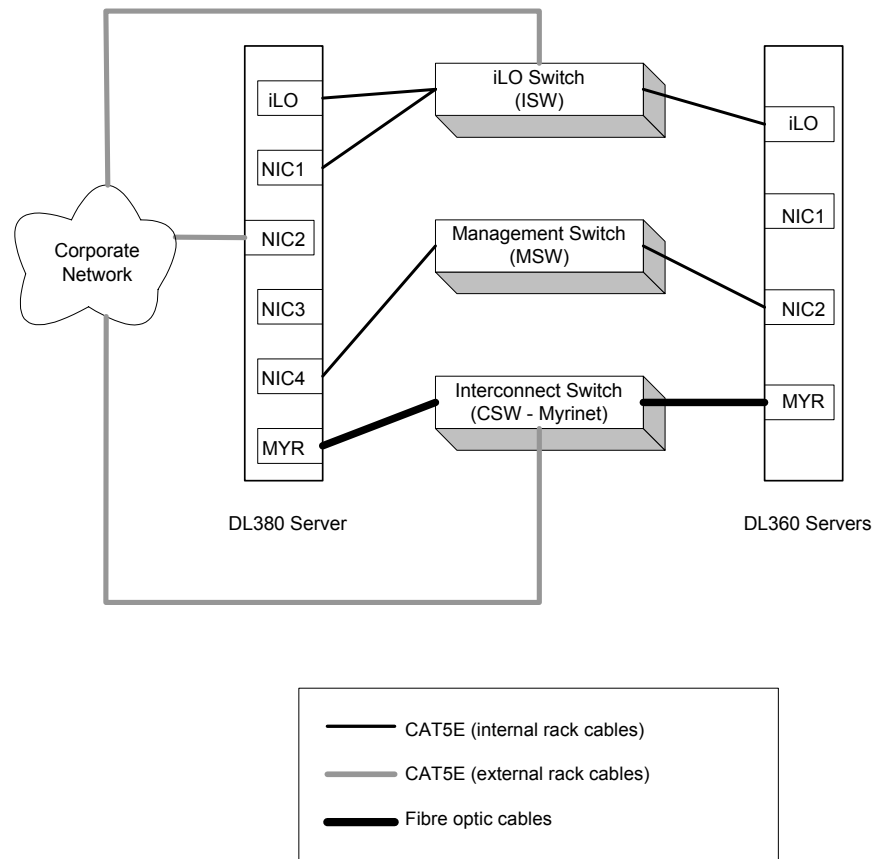

figure 3. network overview of Myrinet HPC cluster solutions

In figure 3 above, the DL380 server has one (1) iLO port, four (4) NIC ports, and one (1) Myrinet adapter connection. Both the iLO port and NIC1 port connect to the iLO switch (ISW). NIC2 is used to connect to an external network. NIC 3 is not connected but may be utilized to connect to an external network. NIC4 is used to connect to the management switch (MSW). The Myrinet adapter connection (MYR) is used to connect to the cluster interconnect switch (CSW). In this configuration, the cluster interconnect switch is a Myrinet switch and may vary in size depending on how may compute nodes the HPC solution contains.

The DL360 servers have one (1) iLO port, two (2) NIC ports, and one (1) Myrinet adapter connection. The iLO port connects to the iLO switch (ISW). NIC1 is not connected. NIC2 is used to connect to the management switch (MSW). The Myrinet adapter connection (MYR) is used to connect to the cluster interconnect switch (CSW). In this configuration, the cluster interconnect switch is a Myrinet switch and may vary in size depending on how may computer nodes the HPC solution contains.

Figure 3 above assumes you will be connecting your HPC LC Series solution to an external DHCP server. The DHCP server must be provided by the customer and be made available to the system before proceeding to use iLO. If you plan on using the control node and HPC cluster software to assign DHCP addresses, then you will need to cable the system accordingly.

powering on the equipment

The equipment should be completely installed into the rack(s) and cabled before powering on any hardware components. First, ensure the PDUs have been plugged into appropriate power receptacles and then powered on. Second, power on the monitor and all switch devices. Ensure all switches have initialized correctly before continuing. Before powering on the servers, please review the next section to learn about the factory system settings for each machine. Then proceed to the following sections for powering on and setting up the servers.

factory system settings

There are several factory system settings that are pre-configured before the HPC cluster is shipped to the customer. The pre-configuration settings of the ProLiant DL380 server and the ProLiant DL360 servers include the following:

ROM-based Setup Utility (RBSU) Settings:
- The operating system setting is set to Linux
- Hyperthreading is disabled

iLO settings:
- The iLO firmware is updated to version 1.40 or later
- The iLO Advanced License is installed on all nodes
- The iLO DNS name is set to match the iLO cable label for each node. For example, if the iLO cable label reads R1-A0-iLO, then the iLO DNS name will be set to R1-A0-iLO for that node.
- The iLO tags for each machine have been removed. The iLO username and password for each machine are pre-defined to facilitate the setup process. The iLO username and password for each machine are as follows:
    o   username = Administrator
    o   password = Administrator

Note: By default, iLO is set to obtain an IP address from a DHCP server. The DHCP server must be provided by the customer and be made available to the systems before proceeding to use iLO for initial setup. For detailed instructions and help in configuring DHCP for your network, please refer to the DHCP – HowTo included with your Linux distribution and on the Linux Documentation Project web site.

RAID settings:
- No RAID settings are configured

setting up the control node

First, power on and set up the ProLiant DL380 server – control node. This server is used as the interface to the user community for job dispatch, control, monitoring, and job completion within the HPC cluster. Please refer to the software vendor's installation instructions for setting up the control node for your specific operating system.

Note: During the installation process you may be required to assign each NIC a specific function within the HPC cluster.

setting up the compute nodes

The ProLiant DL360 servers – compute nodes can be powered on in any order. The compute nodes are not connected to a keyboard, monitor, or public LAN. Therefore, you must use the integrated Lights-Out (iLO) feature of the ProLiant servers in order to remotely manage each compute node. Each ProLiant machine has the iLO Advanced license installed. The iLO Advanced license (or feature suite) offers sophisticated virtual administration features for full control of servers in dynamic data center and remote locations. The iLO Advanced feature suite includes Virtual Graphical Console and Virtual Floppy Drive that provide significant cost savings by removing any advantages of being physically present in front of the server for routine access and maintenance. Another feature of iLO Advanced allows a local client CDROM to be connected to a remote host server as a USB device, removing the need to visit the host server to insert and use a CDROM device.

The control node can be used to establish a remote iLO session to each compute node. Furthermore, you can connect additional iLO client machines to the iLO switch or you can connect the iLO switch to a corporate network if you want more than one iLO client machine.

Note: Default iLO settings have been configured in the factory as specified above. Upon receipt of the HPC cluster solution, the customer may choose to change these default settings. Please refer to the iLO User Guide for details on changing any of these default settings.

Please refer to the software vendor's installation instructions for setting up the compute nodes for your specific operating system.

Note: During the installation process you may be required to assign each NIC a specific function within the HPC cluster.

summary

The HP High Performance Clusters LC Series is a range of pre-configured hardware configurations that are integrated, tested, and shipped assembled in racks ready for customer use. This paper has covered the overall setup and installation process for the HP High Performance Clusters LC Series solutions. This guide provided the customer with the information required to take the product from delivery of order to a fully cabled system, ready for operating system and application installation.

for more information

To learn more about HP High Availability and ProLiant Clusters visit the following Web site: http://www.hp.com/servers/proliant/highavailability.

To learn more about HP High Performance Computing visit the following Web sites:
http://www.hp.com/techservers
http://www.hp.com/techservers/clusters/
http://www.hp.com/techservers/resources/example_customers.html
http://www.hp.com/techservers/support/developers.html

feedback

Help us improve our technical communication. Let us know what you think about the technical information in this document. Your feedback is valuable and helps us structure future communications. Please send your comments to: hawebserver@hp.com